Números de Catalan
Degustaciones Matemáticas

Ernesto Vallejo
Centro de Ciencias Matemáticas, UNAM
vallejo@matmor.unam.mx

Los números de Catalan forman una sucesión de números muy importante en Combinatoria Enumerativa, ya que emergen en una amplia gama de problemas de conteo (véase por Ejemplo [3, 4, 5]). La primera aparición de la que se tiene registro (todavía sin ningún significado combinatorio) se da en la primera mitad del siglo XVIII en un trabajo del matemático mongol Ming Antu (c. 1692-c. 1763) sobre una expansión en serie de potencias de $\operatorname{sen}(2\theta)$ como función de $\operatorname{sen}(\theta)$ (véase [1, 6]). En Europa aparecen poco después en la solución del problema propuesto por Leonhard Euler (1707-1783) en 1751 que consiste en calcular el número C_n de triangulaciones de un polígono convexo con $n+2$ lados. Euler, con ayuda de Christian Goldbach (1690 - 1764) y Johann Andreas von Segner (1704-1777), encontró finalmente la fórmula (2.3). El nombre actual de estos números se deriva del apellido del matemático belga Eugène Charles Catalan (1814-1894), quien descubrió la fórmula (2.1) que actualmente se usa como definición en muchos textos de combinatoria (para más información histórica véase [2]).

1 Coeficientes binomiales

En este escrito usaré la siguiente notación: $\mathbb{N}$ denota el conjunto de los números naturales $\{1, 2, 3, \dots\}$ y $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Para cada $n \in \mathbb{N}_0$, $[\,n\,] = \{i \in \mathbb{N} \mid 1 \le i \le n\}$. En particular, $[\,0\,] = \varnothing$, el conjunto vacío. El ***factorial*** de n se define como $n! = n(n-1)\cdots 2 \cdot 1$ si $n > 0$ y $0! = 1$. La ***cardinalidad*** de un conjunto finito S, esto es, el número de sus elementos, se denotará mediante $\#S$ o $|S|$. Para cada conjunto finito X y cada $k \in \mathbb{N}_0$ definimos el conjunto

$$\binom{X}{k} = \{S \subseteq X \mid \#S = k\}.$$

Si X y Y son dos conjuntos finitos con la misma cardinalidad se cumple $\#\binom{X}{k} = \#\binom{Y}{k}$. Este número, al depender solo de la cardinalidad de X, se denotará como $\binom{|X|}{k}$. Se llama ***coeficiente binomial***. Nótese que para cada $n \in \mathbb{N}_0$ se tiene que $\binom{n}{0} = 1$ y si $k > n$, $\binom{n}{k} = 0$. Aún sin haber calculado explícitamente los números $\binom{n}{k}$ podemos obtener algunas identidades entre ellos usando argumentos puramente combinatorios.

Lema 1.1. *Para cada k, $n \in \mathbb{N}$, tales que $n \ge k \ge 1$, se cumple la **recurrencia de Pascal***

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

Demostración. Sean $X = \{x_1, \dots, x_n\}$ y $\overline{X} = \{x_1, \dots, x_{n-1}\}$. Para cada $S \subseteq X$, denotamos $\overline{S} = S \setminus \{x_n\}$. Entonces, la correspondencia $S \longmapsto \overline{S}$ define una función biyectiva

$$\Phi_k : \binom{X}{k} \longrightarrow \binom{\overline{X}}{k-1} \bigcup \binom{\overline{X}}{k}.$$

Por lo tanto el dominio y el codominio de Φ_k tienen la misma cardinalidad. Dado que la intersección $\binom{\overline{X}}{k-1} \cap \binom{\overline{X}}{k}$ es vacía, se tiene que

$$\#\left[\binom{\overline{X}}{k-1} \bigcup \binom{\overline{X}}{k}\right] = \#\binom{\overline{X}}{k-1} + \#\binom{\overline{X}}{k}.$$

La igualdad queda demostrada. $\square$

Ejercicio 1.1. Demuestra que para cada m, n, $r \in \mathbb{N}_0$ se cumple la ***fórmula de convolución de Vandermonde***

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}.$$

Damos ahora una fórmula explícita para los coeficientes binomiales.

Lema 1.2. *Sean k, $n \in \mathbb{N}$, $k \leq n$. Entonces*

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!\,(n-k)!}.$$

Demostración. Nótese que el número de funciones inyectivas $f : [\,k\,] \longrightarrow [\,n\,]$, $k \leq n$, es igual a $n(n-1)\cdots(n-k+1)$. Puesto que en un conjunto no importa el orden de sus elementos, cada subconjunto de $[\,n\,]$ con k elementos es la imagen de $k!$ funciones inyectivas $f : [\,k\,] \longrightarrow [\,n\,]$. De estas dos observaciones se sigue la primera igualdad. La segunda es una consecuencia directa de la primera. $\square$

2 Números de Catalan

Sea $n \in \mathbb{N}_0$. El n-ésimo ***número de Catalan*** se define como

$$C_n = \frac{1}{n+1}\binom{2n}{n}. \tag{2.1}$$

Los primeros números son: $C_0 = 1$, $C_1 = 1$, $C_2 = 2$, $C_3 = 5$, $C_4 = 14$, $C_5 = 42$, $C_6 = 132$, $C_7 = 429$, $C_8 = 1430$, $C_9 = 4862$ y $C_{10} = 16796$.

Se deja como ejercicio la demostración de algunas identidades de los números de Catalan.

Ejercicio 2.1. Para $n \geq 1$, demuestra que

$$C_n = \binom{2n}{n} - \binom{2n}{n+1}. \tag{2.2}$$

Por definición los números de Catalan son positivos, pero no es evidente que sean enteros; de la identidad (2.2) se sigue que son en realidad enteros positivos.

Ejercicio 2.2. Usa la fórmula de convolución de Vandermonde para demostrar que

$$C_n = \frac{1}{n+1}\sum_{k=0}^{n}\binom{n}{k}^2.$$

Ejercicio 2.3. Demuestra que, para $n \geq 2$,

$$C_n = \prod_{k=2}^{n}\frac{n+k}{k}.$$

Ejercicio 2.4. Demuestra que para $n \in \mathbb{N}$

$$C_n = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{(n+1)!} 2^n.$$

Ejercicio 2.5. Demuestra que para $n \geq 0$ vale

$$\frac{C_{n+1}}{C_n} = \frac{4n+2}{n+2}.$$

Deduce, para $n \geq 1$, la siguiente fórmula debida a Euler (véase [2, p. 178]):

$$C_n = \prod_{k=2}^{n+1} \frac{4k-6}{k}. \tag{2.3}$$

3 Palabras de Dyck

En esta sección mostramos que C_n es el número de palabras de Dyck de longitud $2n$. Esta es una de las muchas interpretaciones combinatorias de los números de Catalan. En [4] el lector interesado puede consultar más de doscientas interpretaciones combinatorias de C_n, entre ellas la que corresponde a la definición de Euler como el número de triangulaciones de un polígono convexo.

Sea $X = \{A, B\}$ un conjunto con dos elementos. Pensemos X como un *alfabeto* y sus elementos como *letras*. Conviene suponer que las letras están ordenadas linealmente, es decir, que $A < B$. Una ***palabra*** en X de ***longitud*** n es una función

$$w : [\,n\,] \longrightarrow X.$$

La escribimos como $w = w_1 \cdots w_n$, donde $w_i = w(i)$. La única palabra de longitud 0 es la ***palabra vacía***, que denotamos como $\varnothing$. Si w es una palabra de longitud n y $k \leq n$, el ***segmento inicial*** de w de longitud k es la restricción de w a $[\,k\,]$ y lo denotamos como $[w]_k$. Así,

$$[w]_k = w_1 \cdots w_k.$$

Sean x in X y w una palabra en el alfabeto X. La ***frecuencia*** de x en w es el número de veces que la letra x aparece en la palabra w:

$$F_x(w) = \# w^{-1}(x).$$

Denotamos mediante $\mathcal{P}_{a,b}$ el conjunto de todas las palabras w en el alfabeto X tales que A aparece en w con frecuencia a y B aparece en w con frecuencia b. El ***complemento*** w^{c} de w se forma reemplazando en w la letra A por B y la letra B por A. Por lo cual,

$$w \in \mathcal{P}_{a,b} \quad \Longleftrightarrow \quad w^{\mathsf{c}} \in \mathcal{P}_{b,a}.$$

Denotamos la diferencia de las frecuencias de A y B es una palabra w como

$$\Delta_{A,B}(w) = F_A(w) - F_B(w).$$

Por ejemplo, si $w = AABBAB$, entonces $w^{\mathsf{c}} = BBAABA$, $F_A([w]_3) = 2$, $F_B([w]_3) = 1$ y $\Delta_{A,B}([w]_3) = 1$. Llegamos, por fin, a los dos conceptos importantes de esta sección:

Definición 3.1. Sea w una palabra en el alfabeto X de longitud positiva n. Decimos que
(1) w es una ***palabra de Dyck fuerte*** si para cada $k \in [\,n\,]$ se cumple $\Delta_{A,B}([w]_k) > 0$;
(2) w es una ***palabra de Dyck débil*** si para cada $k \in [\,n\,]$ se cumple $\Delta_{A,B}([w]_k) \geq 0$.

Lema 3.2. *Sean a, $b \in \mathbb{N}$. Entonces*

$$|\mathcal{P}_{a,b}| = \binom{a+b}{a}.$$

Demostración. Sea $n = a + b$. Luego, si $w \in \mathcal{P}_{a,b}$, w tiene longitud n. Sea $S_{n,a}$ el conjunto de subconjuntos de $[\,n\,]$ con a elementos. Por lo tanto, $|S_{n,a}| = \binom{n}{a}$. Para demostrar nuestra afirmación bastará construir una función biyectiva entre $\mathcal{P}_{a,b}$ y $S_{n,a}$. Dada $w \in \mathcal{P}_{a,b}$, definimos $T_w = \{i \in [\,n\,] \mid w_i = a\}$. La correspondencia $w \longmapsto T_w$ define una función $T : \mathcal{P}_{a,b} \longrightarrow S_{n,a}$. Es rutina el verificar que T es biyectiva. Con esto el lema queda demostrado. □

Teorema 3.3. *Sean a, $b \in \mathbb{N}$ tales que $a > b$. El número de palabras de Dyck fuertes en $\mathcal{P}_{a,b}$ es igual a*

$$\frac{a-b}{a+b}\binom{a+b}{a}.$$

Proof. Cualquier palabra de Dyck fuerte empieza necesariamente con una A. Así, consideramos el conjunto

$$\mathcal{P}_{a,b}(A) = \{w \in \mathcal{P}_{a,b} \mid w_1 = A\}.$$

Notemos que, por el Lema 3.2,

$$|\mathcal{P}_{a,b}(A)| = \binom{a+b-1}{a-1}. \tag{3.1}$$

Resulta más fácil contar los elementos en $\mathcal{P}_{a,b}(A)$ que no son palabras de Dyck fuertes, esto es, aquellas palabras para las que hay una s tal que $\Delta_{A,B}([w]_s) \leq 0$. Si w no es una palabra de Dyck fuerte, como $\Delta_{A,B}(w) = a - b > 0$, existe t tal que $\Delta_{A,B}([w]_t) = 0$. Sea

$$T = \{w \in \mathcal{P}_{a,b}(A) \mid \text{existe } t \in [\,a+b\,] \text{ que satisface } \Delta_{A,B}([w]_t) = 0\}.$$

Entonces un elemento $w \in \mathcal{P}_{a,b}(A)$ es palabra de Dyck fuerte si y solo si $w \notin T$. Sea $D^f_{a,b}$ el número de palabras de Dyck fuertes en $\mathcal{P}_{a,b}$. Por lo tanto,

$$D^f_{a,b} = |\mathcal{P}_{a,b}(A)| - |T|. \tag{3.2}$$

Para calcular $|T|$ definimos una función biyectiva

$$\theta : T \longrightarrow \mathcal{P}_{b,a}(A).$$

Sea $w \in T$. Sea $k = \max\{t \in [\,a+b\,] \mid \Delta_{A,B}([w]_t) = 0\}$. Luego, $1 < k < a + b$. Sea z la palabra tal que $w = [w]_k z$. Definimos $\theta(w) = [w]_k z^{\mathsf{c}}$. Como k es par, lo podemos escribir como $k = 2m$. Entonces $F_A(z) = a - m$ y $F_B(z) = b - m$. Por lo tanto $F_A(\theta(w)) = m + (b - m) = b$ y $F_B(\theta(w)) = m + (a - m) = a$. Así, θ es una función bien definida.

Demostramos ahora que θ es inyectiva. Sean v, $w \in T$ tales que $\theta(v) = \theta(w)$. Supongamos que $\theta(v) = [v]_k x^{\mathsf{c}}$ y $\theta(w) = [w]_\ell y^{\mathsf{c}}$. Como k es el último empate entre A's y B's en v, k es también el último empate entre A's y B's en $\theta(v)$. Esto fuerza que $k = \ell$. Luego, $x^{\mathsf{c}} = y^{\mathsf{c}}$. Por lo cual $x = y$ y $v = w$.

Resta demostrar que θ es suprayectiva. Sea $v \in \mathcal{P}_{b,a}(A)$. Por lo tanto, v empieza con A, $F_A(v) = b$ y $F_B(v) = a$. Dado que $a > b$, existe r tal que $\Delta_{A,B}([v]_r) = 0$. Sea ℓ el máximo entre todas las t's tales que $\Delta_{A,B}([v]_t) = 0$. Sea y la palabra tal que $v = [v]_\ell y$. Entonces $[v]_\ell y^{\mathsf{c}} \in T$ y $\theta([v]_\ell y^{\mathsf{c}}) = v$.

Hemos demostrado que θ es biyectiva, por lo cual $|T| = |\mathcal{P}_{b,a}(A)|$. Así, de la ecuación (3.2) se obtiene que

$$D^f_{a,b} = |\mathcal{P}_{a,b}(A)| - |\mathcal{P}_{b,a}(A)|.$$

Al aplicar la igualdad (3.1) se llega a que

$$\begin{aligned} D^f_{a,b} &= \binom{a+b-1}{a-1} - \binom{a+b-1}{b-1} = \frac{(a+b-1)!}{(a-1)!\,b!} - \frac{(a+b-1)!}{a!\,(b-1)!} \\ &= \frac{(a-b)(a+b-1)!}{a!\,b!} = \frac{a-b}{a+b}\binom{a+b}{a}. \end{aligned}$$

El teorema queda demostrado. □

Ahora podemos obtener fácilmente el número de palabras de Dyck débiles.

Corolario 3.4. *Sean a, $b \in \mathbb{N}$ tales que $a \geq b$. El número de palabras de Dyck débiles en $\mathcal{P}_{a,b}$ es igual a*

$$\frac{a+1-b}{a+1}\binom{a+b}{a}.$$

Demostración. Sea $w \in \mathcal{P}_{a,b}$. Dado que w es una palabra de Dyck débil si y solo si Aw es una palabra de Dyck fuerte, se sigue del Teorema 3.3 y del Lema 1.2 que el número de palabras de Dyck débiles es igual a

$$\frac{a+1-b}{a+1+b}\binom{a+1+b}{a+1} = \frac{a+1-b}{a+1}\binom{a+b}{a}.$$

La identidad queda demostrada. □

Un caso particular del corolario anterior nos da una interpretación combinatoria de los números de Catalan. Una ***palabra de Dyck*** de longitud $2n$ en el alfabeto X es una palabra de Dyck débil en $\mathcal{P}_{n,n}$. Consideramos la palabra vacía $\varnothing$ como una (la única) palabra de Dyck de longitud 0.

Teorema 3.5. *Sea $n \in \mathbb{N}_0$. Entonces, el número de palabras de Dyck de longitud $2n$ es C_n.*

Demostración. Es una consecuencia inmediata del Corolario 3.4 cuando $a = b = n$ y de la definición de número de Catalan dada en (2.1). □

Una palabra de Dyck se puede representar geométricamete mediante un camino en $\mathbb{Z}^2$ llamado ***camino de Dyck***. Esta representación puede resultar más intuitiva a algunos los lectores (véase, por Ejemplo, [3, p. 11] o [5]).

References

[1] P. J. Larcombe, *The 18th Century Chinese Discovery of the Catalan Numbers*, Mathematical Spectrum, **32** No. 1, 1999 / 2000, 5–7.

[2] I. Pak, *History of Catalan Numbers*, Apéndice B en *Catalan Numbers*, R. P. Stanley, Cambridge University Press, Nueva York, 2015.

[3] S. Roman, *An Introduction to Catalan Numbers*, Birkhäuser, Heidelberg, 2015.

[4] R. P. Stanley, *Catalan Numbers*, Cambridge University Press, Nueva York, 2015.

[5] Wikipedia contributors. *Catalan number*. Wikipedia, The Free Encyclopedia, 7 May. 2019. Internet, descargado 18-junio-2019.

[6] Wikipedia contributors. *Minggatu*. Wikipedia, The Free Encyclopedia, 1 Apr. 2019. Internet, descargado 18-junio-2019.